reprint
from
Department of Educational and Psychological Research
Malmö School of Education - University of Lund
No. 284

EIN NEUES VERFAHREN ZUR LÖSUNG PSYCHOMETRISCHER PROBLEME IN DER ANALYSE VERBALER DATEN

Bernhard Bierschenk

In: Eckensberger, L. H. (Hrsg.) **Bericht über den 31. Kongress der Deutschen Gesellschaft für Psychologie.** Göttingen: Hogrefe, 1979. Ss. 80-84.

Sonderdruck aus:
Bericht über den 31. Kongreß der Deutschen Gesellschaft für Psychologie

EIN NEUES VERFAHREN ZUR LÖSUNG PSYCHOMETRISCHER PROBLEME IN DER ANALYSE VERBALER DATEN

Bernhard Bierschenk
Department of Educational
and Psychological Research
University of Lund-Malmö

Verbale Äusserungen werden in der verhaltenswissenschaftlichen Forschung mit den unterschiedlichsten Methoden und Techniken analysiert. Um die Relationen innerhalb eines ganzen Satzes quantitativ beschreiben und mit Hilfe multivariater statistischer Verfahren analysieren zu können, wurde ANACONDA entwickelt. Diese Abkürzung ist die Bezeichnung einer Methode zur "Analysis of Concepts by means of Data Processing".

Der Methode liegt ein psychologisches Modell zugrunde, das von der Annahme ausgeht, dass eine empirische Beobachtung aus einem ganzen Satz besteht. Ausserdem wird angenommen, dass jeder verbalen Äusserung kognitive Einheiten ("Begriffe") zugrunde liegen, die sich in den Schlüsselwörtern eines Satzes manifestieren. Weiterhin wird angenommen, dass ein Satz nicht etwa das Produkt strikt zufälliger Prozesses ist, sondern vielmehr stochastische Prozesse voraussetzt.

Der Entwicklung von ANACONDA liegt ein empirisches Material zugrunde, das aus Interviewtexten besteht. Diese wurden mit Hilfe von Klusteranalysen kondensiert und homogenisiert. Die Skalierung einzelner linguistischer Elemente wurde mit Hilfe eines Panels durchgeführt. An dem Panel haben 15 Verhaltenswissenschaftler der Universität Lund-Malmö (Schweden) teilgenommen. Das Relationsmuster in den erhaltenen Kovariationsschemata wurde mit Hilfe einer Diskriminanzanalys untersucht. Schliesslich wurde versucht, die Validität der Methode zu demonstrieren (Bierschenk, 1977).

1. Voraussetzungen

Der Aufbau von Begriffen in einem gegebenen Kontext setzt ein Regelsystem voraus. Damit wird bestimmt, auf welche Weise einzelne linguistische Elemente miteinander verknüpft werden sollen und in welcher Folge einzelne Verknüpfungen geschehen sollen (vgl. Bierschenk & Bierschenk, 1976; B. Bierschenk, 1977). ANACONDA setzt den Aufbau empirisch verankerter Begriffe voraus. Als Bausteine dienen dabei linguistische Elemente wie z.B. Verb oder Adjektiv, während die Grundeinheit der Analyse mit dem Ausgangspunkt in dem wohlbekannten Agent-Aktion-Objekt(Ziel)-Paradigma beschrieben wird. Die Komponenten wurden wie folgt definiert:

Agent: Aktionszentrum oder zielsuchende Entität, welche in der Realisierung ihrer Ziele Hilfsmittel anwenden. (Diese Definition umfasst auch Gruppen, Organisationen oder Abstraktionen, die die Funktion eines Agenten wahrnehmen.)

Aktion: Eine gerichtete Handlung, die mit der Absicht ausgeführt wird, ein Ziel zu realisieren.

Objekt (Ziel): Alles, auf was eine Handlung gerichtet oder gegenüber welchem eine Handlung ausgeführt werden kann.

Durch das AaO-Paradigma werden die Komponenten abgegrenzt, die einen natürlichen Zusammenhang bilden. Damit wird im Hinblick auf eine verbale Aussage die Substantiv_1-Verb-Substantiv_2-Relation, d.h. eine beobachtbare Aussage, gemeint.

ANACONDA baut u.a. auf der Annahme auf, dass die zwischen Substantiv und Adjektiv sowie zwischen Substantiv und Verb bestehenden Relationen solche Beziehungen widerspiegeln, die empirische Phänomene miteinander verknüpfen. Auf dem Hintergrund dieser Annahme wurden sämtliche Verben und Adjektive eines umfassenden Interviewmaterials skaliert. Es kann also einer Gruppe von Substantiven durch eine Gruppe von Adjektiven ein empirischer und quantitativ definierter Inhalt gegeben werden, d.h. es können quantitativ definierte Begriffe aufgebaut werden.

Weiterhin wird angenommen, dass der Inhalt, der der Sprache eines einzelnen und den sich daraus ergebenden Transformationen zugrunde liegt, einzig und allein nur von empirischer Natur sein kann. Gemäss Quine (1972) kann daher nur ein vollständiger Satz die Grundeinheit einer Analyse und Synthese empirischer Phänomene sein. Die charakteristische Eigenschaft eines solchen Satzes ist dessen Intersubjektivität. Rozeboom (1972, S 97) behauptet ebenfalls, dass Wissen nichts anderes sein kann als "propositional knowledge" bzw. "justified true belief".

Die linguistische Grundeinheit in ANACONDA ist daher das bereits erwähnte Substantiv_1-Verb-Substantiv_2-Paradigma. Wird diesem eine psychologische Dimension hinzugefügt, ergibt sich das Agent-Aktion-Objekt-Paradigma. Der wesentliche Unterschied zwischen beiden ist, dass im zweiten eine Richtungsindikation gefordert wird. Dieses Paradigma, verkürzt mit AaO, fordert also eine explizite Angabe des Handelnden bzw. des Agenten (A), des Ereignisses, bzw. der Aktion (a) und des Objektes, bzw. des Ziels (O) für die Handlung. Entsprechend des AaO-Paradigmas ist eine Handlung immer zielend (transitiv). Dabei spielt es keine Rolle, ob diese explizit oder implizit angegeben wird.

Es kann somit bereits zwischen zwei Ebenen im Analysenverfahren unterschieden werden, nämlich

1. einer manifesten Ebene, auf der die Repräsentationsform sprachlicher Produktionen von entscheidender Bedeutung ist, und
2. einer latenten Ebene, auf der Modelle zur Approximation von Phänomenen psychologischer Natur die entscheidene Rolle spielen.

Auf der manifesten Ebene hängt die Qualität des Analysenverfahrens entscheidend von der Kontrolle über das Verb ab. Durch das Verb wird nämlich die Bedeutung des Substantiv_1-Verb-Substantiv_2-Paradigmas, verkürzt mit $S_1v\ S_2$, angegeben und die einzelnen Komponenten des Paradigmas identifiziert, d.h. ob eine vollständige Aussage vorliegt oder nicht.

Auf der latenten Ebene hängt die Qualität des Analysenverfahrens von den theoretischen Begriffsdefinitionen ab. Auf dieser Ebene werden die einzelnen Komponenten im AaO-Paradigma durch die Handlung identifiziert, d.h. die Handlung definiert den Inhalt. Kann angenommen werden, dass zwischen den generierten und quantitativ definierten Begriffen Relationen implikativen Charakters bestehen, d.h. Kovariationen zwischen den Eigenschaften, kann auch die Kognition von Objekten vorhergesagt werden.

Durch eine Skalierung von Adjektiven und Verben und deren Kopplung mit ihren jeweiligen Substantiven wird es möglich, Texte numerisch zu beschreiben und multivariaten Analysen zu unterziehen. Ein solches Verfahren erlaubt es auch, die Position eines Begriffes simultan in einer Anzahl latenter Dimensionen zu bestimmen.

ANACONDA nimmt Rücksicht auf "syntaktische Verhältnisse" und betrachtet sowohl Adjektive als auch Verben als abhängige, bzw. als beschreibende linguistische Elemente. Substantive werden dagegen als unabhängige, bzw. als zu beschreibende linguistische Elemente betrachtet. Adjektive beschreiben Substantive direkt, während Verben die gleiche Funktion indirekt erfüllen.

2. Adjektiv- und Verbskalierung

Eine Methode für die Bestimmung der Intersubjektivität in Beurteilungen ist die Panelbeurteilung. Mit Hilfe von siebenstufigen Skalen wurden Panelbeurteilungen für 570 Adjektive und 883 Verben durchgeführt.

Um verschiedenen Fehlerquellen entgegen zu wirken, wurden für jeden einzelnen Beurteiler individuelle und strikt zufällig bestimmte Wortlisten hergestellt. Die abgegebenen Beurteilungen (insgesamt 8 718 Werte) wurden strukturanalytisch untersucht. Die Reliabilitätsschätzungen für die einzelnen Skalen wurden mit einem von Lord (1958) vorgeschlagenen Verfahren zur Berechung eines Koeffizienten für maximale Reliabilität (α_{max}) durchgeführt. In Tabelle 1 sollen die Reliabilitätsschätzungen in sehr kondensierter Form reproduziert werden.

Tabelle 1. Reliabilitäten in den Adjektiv- und Verbbeurteilungen (α_{max})

Beurteilt	Dimension (1) Wertung	(2) Aktivität	(3) Stärke
Adjektiv	.97	.92	.88
Verb	.95	.93	.86

Eine detaillierte Darstellung der Reliabilitätsuntersuchung und der Diskussion der erhaltenen Resultate wird in Bierschenk & Bierschenk (1976) berichtet.

3. Begriffsgenerierung

ANACONDA impliziert eine Analyse und Synthese von sowohl empirischen Aussagen als auch den Relationen zwischen den Aussagen. Aus diesem Grund reicht

eine lexikographische Aufstellung von Wörtern und die Konstruktion eines Kategoriensystems nicht aus, wenn es gilt, die impliziten kognitiven Modelle einer Informationsquelle zu erfassen. Um jedoch diesem Ziel näher kommen zu können, wurden im Grundmaterial verankerte Wortlisten geschaffen und auf der Basis der skalierten Elemente quantitativ bestimmte Begriffe generiert. Dem Generierungsprozess liegen sowohl horizontelle als auch vertikale Relationen innerhalb einer Aussage zugrunde.

Das Analysenverfahren wurde in sechszehn Schritten durchgeführt. Die Schritte eins bis vier des Verfahrens wurden durchgeführt, um untersuchen zu können, in welchem Ausmass gemeinsame linguistische Elemente in den verbalen Formulierungen angewendet werden. Die resultierenden Frequenzverteilungen bildeten den Ausgangspunkt für eine Kondensierung des Materials durch Klusteranalysen. In den Schritten fünf bis sieben wurden verschiedene Kriterien für Klusterbildungen und Zusammenschlagungen geprüft. In den Schritten acht bis zwölf wurden einzelne Gruppierungen auf untere Grenzwerte hin untersucht. Im dreizehnten und vierzehnten Schritt werden die A- und O-Komponenten aufs neue geklustert. Im fünfzehnten Schritt werden die einzelnen Kluster auf ihre Koinzidenzen hin untersucht. Im sechzehnten Schritt des Verfahrens kommen schliesslich die Skalierungswerte der Verben zur Anwendung, d.h. die Objekte werden quantifiziert und damit zu empirisch verankerten Begriffen. Diese werden dann auf ihre Diskrimination im Hinblick auf (1) Evaluation, (2) Aktivität und (3) Stärke hin untersucht.

Gewöhnlich wird eine Relation damit zum Ausdruck gebracht, dass eine Regel formuliert wird, die besagt, was als Element, Paar usw. betrachtet werden soll. Wird für eine Stichprobe (S) angegeben, dass A eine Teilmenge der Menge aller Nomina und B eine Teilmenge aller Adjektive ist, die A modifizieren, kann dieses Verhältnis formalisiert und zur quantitativen Bestimmung von Begriffen, d.h. den Komponenten im AaO-Paradigma, benutzt werden. Eine statistisch sinnvolle Analyse und Synthese fordert aber gleichzeitig Computerprogramme, die eine Verarbeitung grosser Mengen verbaler Daten ermöglichen. Computerprogramme wurden zu diesem Zweck entwickelt.

4. Latente Strukturen im Interviewtext

Die Resultate aus den Klusteranalysen zusammen mit den Skalierungsresultaten lassen eine Diskriminanzanalyse zu. Die genannten drei Dimensionen in den Schätzungen können als drei Gruppierungen der Schätzungswerte aufgefasst, d.h. es können zwei Diskriminanzfunktionen gebildet werden. Der Analyse liegen die Kluster der Agenten als Objekte der Messung zugrunde. Die Kluster der Objekte werden dagegen als Variablen betrachtet. Die hergeleiteten Diskriminanzfunktionen (f) werden in Tabelle 2 dargestellt.

Tabelle 2. Diskriminanzfunktionen und standardizierte Koeffizienten

Begriffskluster	Funktion f_1	f_2
Bibliographische Referenz	.50	.04
Problemdiskussion	.44	.82
Informationskanal	.53	-.71
Forschungsmethodische Information	-.17	-.49
Begriffliche Abgrenzung	.32	.33

Die Koeffizienten in Tabelle 2 sind Gewichte und können auf gleiche Weise wie Betakoeffizienten in einer multiplen Regressionsanalyse interpretiert werden. Für die erste Funktion sind "Bibliographische Referenz" und "Informationskanal" von Bedeutung. Das Kluster "Informationskanal" zeigt ausserdem ein hohes negatives Gewicht in der zweiten Funktion. Für die zweite Funktion sind "Problemdiskussion" und mit negativem Vorzeichen "Forschungsmethodische Information" von Bedeutung. "Begriffsmässige Abgrenzung" ist dagegen für beide Funktionen von gleicher Bedeutung.

Diese Verhältnisse können näher untersucht werden, indem die Kluster der Agenten im Hinblick auf die von B. Bierschenk (1977) berichteten Schätzungswerte untersucht werden. An der Relation zwischen "Informationskanal" und "Bibliographische Referenz" soll abschliessend ein solches Vorgehen veranschaulicht werden:

A	a	O
Informationskanal		Bibliographische Referenz
mit den Agenten		mit den Objekten
Bibliotek		Literatur
Psychological Abstracts		Referenz
		Zeitschrift

Die beide verknüpfenden Handlungen geben Ausdruck für eine positive Wertung ($m = 4.60$), eine gewisse Aktivität ($m = 4.21$) und eine geringe Stärke ($m = 3.76$).

Literatur

Bierschenk, B. A computer-based content analysis of interview texts: Numeric description und multivariate analysis. Didakometry, No 53, 1977.

Bierschenk, B. & Bierschenk, I. A system for a computer-based content analysis of interview data. (Studia Psychologica et Paedagogica, 32.) Lund: Gleerup, 1976.

Lord, F.M. Some relations between Guttman's principal components of scale analysis and other psychometric theory. Psychometrika, 1958, 23, 291-296.

Quine, W.V. Epistemology naturalized. In: J.R. Royce & Wm.W. Rozeboom. (Eds.) The psychology of knowing. New York: Gordon & Breach, 1972. Pp. 9-23.

Rozeboom, Wm.W. Problems in the psycho-philosophy of knowledge. In: J.R. Royce & Wm.W. Rozeboom. (Eds.) The psychology of knowing. New York: Gordon & Breach, 1972. Pp. 25-109.